JOURNAL OF CLINICAL MICROBIOLOGY, Feb. 1993, p. 406–409
0095-1137/93/020406-04$02.00/0

Vol. 31, No. 2

NOTES

Strain Identification of *Mycobacterium tuberculosis* by DNA Fingerprinting: Recommendations for a Standardized Methodology

JAN D. A. VAN EMBDEN,[1]* M. DONALD CAVE,[2] JACK T. CRAWFORD,[3] JEREMY W. DALE,[4] KATHLEEN D. EISENACH,[2] BRIGITTE GICQUEL,[5] PETER HERMANS,[6] CARLOS MARTIN,[7] RUTH McADAM,[8] THOMAS M. SHINNICK,[3] AND PETER M. SMALL[9]

Unit Molecular Microbiology, National Institute of Public Health and Environmental Protection, P.O. Box 1, 3720 BA Bilthoven, The Netherlands[1]; University of Arkansas for Medical Sciences, Little Rock, Arkansas 72205[2]; Division of Bacterial and Mycotic Diseases, National Center for Infectious Diseases, Centers for Disease Control, Atlanta, Georgia 30333[3]; Department of Microbiology, University of Surrey, Guildford GU2 5XH, United Kingdom[4]; Microbial Engineering Unit, Institut Pasteur, 75724 Paris, Cedex 15, France[5]; Armauer Hansen Research Institute, Addis Ababa, Ethiopia[6]; Area Microbiologia, Faculty of Medicine, University of Zaragoza, Zaragoza, Spain[7]; Albert Einstein College of Medicine, Yeshiva University, Bronx, New York 10461[8]; and Beckman Center B251, Howard Hughes Medical Institute, Stanford University, Stanford, California 94305-5425[9]

Received 26 August 1992/Accepted 26 October 1992

DNA fingerprinting of *Mycobacterium tuberculosis* has been shown to be a powerful epidemiologic tool. We propose a standardized technique which exploits variability in both the number and genomic position of IS*6110* to generate strain-specific patterns. General use of this technique will permit comparison of results between different laboratories. Such comparisons will facilitate investigations into the international transmission of tuberculosis and may identify specific strains with unique properties such as high infectivity, virulence, or drug resistance.

Epidemiologic studies of tuberculosis can be greatly facilitated by the application of strain-specific markers. Unusual antibiotic susceptibility patterns and phage typing, which have been used for this purpose, have significant limitations. The recently discovered transposable elements in *Mycobacterium tuberculosis* have been shown to be of great potential for use in strain differentiation. *M. tuberculosis* strain typing has already proven to be extremely useful in outbreak investigations (2, 7, 13) and is being applied to a variety of epidemiologic questions in numerous laboratories.

The existence of repetitive DNA elements in *M. tuberculosis* and their potential for use in fingerprinting of *M. tuberculosis* isolates was recognized independently by Eisenach et al. (4), Zainuddin (14), and Zainuddin and Dale (15). The sequence of one of these elements, designated IS*6110*, was first reported by Thierry et al. (11, 12) and was shown to be related to the IS*3* family of insertion sequences which were discovered in members of the family *Enterobacteriaceae*. McAdam et al. (9) independently sequenced the element isolated by Zainuddin and Dale (15), which was designated IS*986*. Subsequently, a related element from *Mycobacterium bovis* BCG was sequenced by Hermans et al. (6) and is referred to as IS*987*. These three sequences differ in only a few base pairs and therefore can be considered essentially the same element. To avoid confusion, we recommend the designation IS*6110* for these elements, except when a specific copy is concerned.

The effectiveness of this insertion sequence (IS) typing system for epidemiological analysis of *M. tuberculosis* isolates has been demonstrated in a number of studies (1–3, 5, 7, 8, 10, 13). In principle, the results obtained by testing large numbers of strains in different laboratories could be compared. This would allow strains from different geographic areas to be compared and the movement of individual strains to be tracked. Such data may provide important insights into the global transmission of tuberculosis and identify strains with particular properties, such as high infectivity, high virulence, and/or multidrug resistance. Analysis of large numbers of isolates may provide answers to long-standing questions regarding the efficacy of BCG vaccination and the frequency of reactivation versus reinfection, which are increasingly important in light of the AIDS pandemic. These large-scale studies will require the use of computer-assisted analysis and comparison of DNA fingerprints. This report describes such a standard method, which will be adopted in our laboratories, and recommends its use to other laboratories so that the results obtained by different laboratories can be compared.

Bacterial growth, DNA extraction, digestion of DNA, and Southern blotting were done as described previously (13). Agarose gels were loaded with a mixture of 1 μg of *Pvu*II-digested genomic *M. tuberculosis* DNA and molecular size marker DNA. *Pvu*II-digested supercoiled ladder DNA (4 ng;

* Corresponding author.

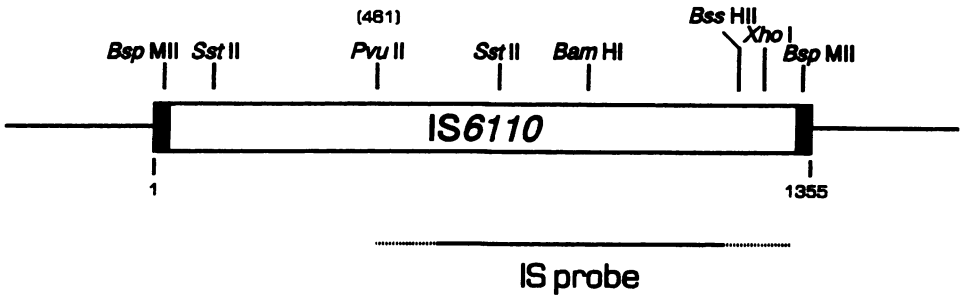

FIG. 1. Physical map of the 1.35-kb *M. tuberculosis* insertion element IS*6110* (9). The cleavage sites of several restriction enzymes are depicted. *Pvu*II cleaves the element at base pair 461. Therefore, any chromosomal mycobacterial DNA fragment obtained by the recommended standard typing method is larger than 0.9 kb. The closed bars represent the 28-bp inverted repeats bordering IS*6110* DNA. The lines to the left and right denote chromosomal DNA.

Bethesda Research Laboratories, Gaithersburg, Md.) and 1.2 ng of *Hae*II-digested φX174 DNA (Clontech, Palo Alto, Calif.) were used as molecular size standards. The molecular sizes of these two reference markers ranged from 16.2 to 0.603 kb, respectively (see Fig. 2B).

The mycobacterial IS probe was prepared by peroxidase labeling of a 245-bp fragment obtained by amplification by the polymerase chain reaction (PCR) described previously (7). Briefly, the oligonucleotides INS-1 (5'-CGTGAGGGCA TCGAGGTGGC) and INS-2 (5'-GCGTAGGCGTCGGTGA CAAA) were used to amplify a 245-bp fragment from purified chromosomal *M. bovis* BCG DNA by PCR. This fragment was purified by Sephadex G50 chromatography. The DNA was precipitated with ethanol, and after solubilization the DNA was labeled with peroxidase as described previously (7).

Standard method of fingerprinting. DNA typing of *M. tuberculosis* complex strains is based on polymorphisms generated by variabilities in both the copy numbers and the chromosomal positions of IS*6110* among clinical isolates of *M. tuberculosis* (1–3, 5–8, 10, 13, 15). The technique of fingerprinting entails the growth of *M. tuberculosis*, extraction of DNA, restriction endonuclease digestion, Southern blotting, and probing for the IS element. Only three parameters are critical for a standardized IS*6110*-based DNA fingerprinting system: the specificity of the restriction enzyme, the nature of the DNA probe, and the use of appropriate molecular mass standards.

The physical map of the IS*6110* sequence (Fig. 1) indicates that various restriction enzymes cleave within the 1,355-bp element. *Bam*HI, *Sst*II, *Pst*I, *Bst*EII, *Bss*HII, and *Pvu*II have all been successfully used to generate restriction fragment length polymorphisms (1–8, 13, 14). For the standard method we recommend the use of *Pvu*II, because it has been used by the majority of laboratories and it cleaves the IS*6110* sequence only once. Because of this latter property, *Pvu*II digestion of IS*6110*-containing genomic DNA leads to IS*6110*-hybridizing fragments of at least 0.90 or 0.46 kb, depending on the IS*6110* probe that is used. Since *M. tuberculosis* usually contains 8 to 20 IS*6110* copies (13), the use of a DNA probe which overlaps both sides of the *Pvu*II site would result in 16 to 40 bands. This large number of bands would result in overcrowded lanes with overlapping bands. Therefore, we arbitrarily chose a DNA probe to the right of the *Pvu*II site on the physical map, as shown in Fig. 1. This reduced the number of IS*6110*-containing bands in the fingerprint to half of the maximum number possible. In exceptional cases, when the differentiation of the patterns is not adequate, the membranes could be reprobed with labeled DNA containing only IS sequences to the left of the *Pvu*II site. For the standardized method, the exact DNA sequences of the probe do not matter as long as the IS sequence to the right of the *Pvu*II site is used. An illustration of a Southern blot is given in Fig. 2A.

In order to compare fingerprints between *M. tuberculosis* isolates run on different gels and in different laboratories, the size of each IS*6110*-hybridizing fragment must be determined. This requires the use of molecular size markers which span the 10- to 0.9-kb range of most IS*6110*-hybridizing fragments. We recommend use of a combination of external and internal standards which provide a compromise between technical ease and maximal precision. External molecular size markers should be run in two or three lanes of each gel. Furthermore, we recommend inclusion in each gel a lane containing DNA from the reference strain of *M. tuberculosis* Mt14323, which, when digested with *Pvu*II and probed with IS*6110*, gives 10 approximately evenly spaced bands of known size (Fig. 2A). Although the use of external markers is adequate for comparing small numbers of similar strains (such as in outbreak investigations), it may not provide sufficient precision to permit computerized comparisons of hundreds or thousands of strains. For this reason we recommend that molecular size standards which do not hybridize with IS*6110* also be added to the wells with the cleaved *M. tuberculosis* DNA. After hybridization with IS*6110*, the membrane can be reprobed with labeled molecular size marker standards. This results in a second autoradiograph with molecular size standards in each lane. The second autoradiograph can be superimposed on the first autoradiograph, resulting in extremely precise molecular size determinations (Fig. 2B). We were able to obtain standard deviations of the molecular sizes of less than 2% in the 1- to 2-kb range and less than 5% in the 9- to 10-kb range (2a). If internal standards are used, a single reference external marker is sufficient.

Finally, to be able to compare DNA fingerprints made in different laboratories, a minimal resolving power of the gels is needed. At a given agarose concentration, the resolving power mainly depends on the electrophoresis time. We recommend conditions such that the distance between the 0.9-kb marker (the approximate size of the smallest *Pvu*II IS-containing fragment) and the 10-kb marker is at least 10 cm.

These recommendations will permit comparison of DNA fingerprints of *M. tuberculosis* made in different laboratories that can use their own optimized procedures for DNA

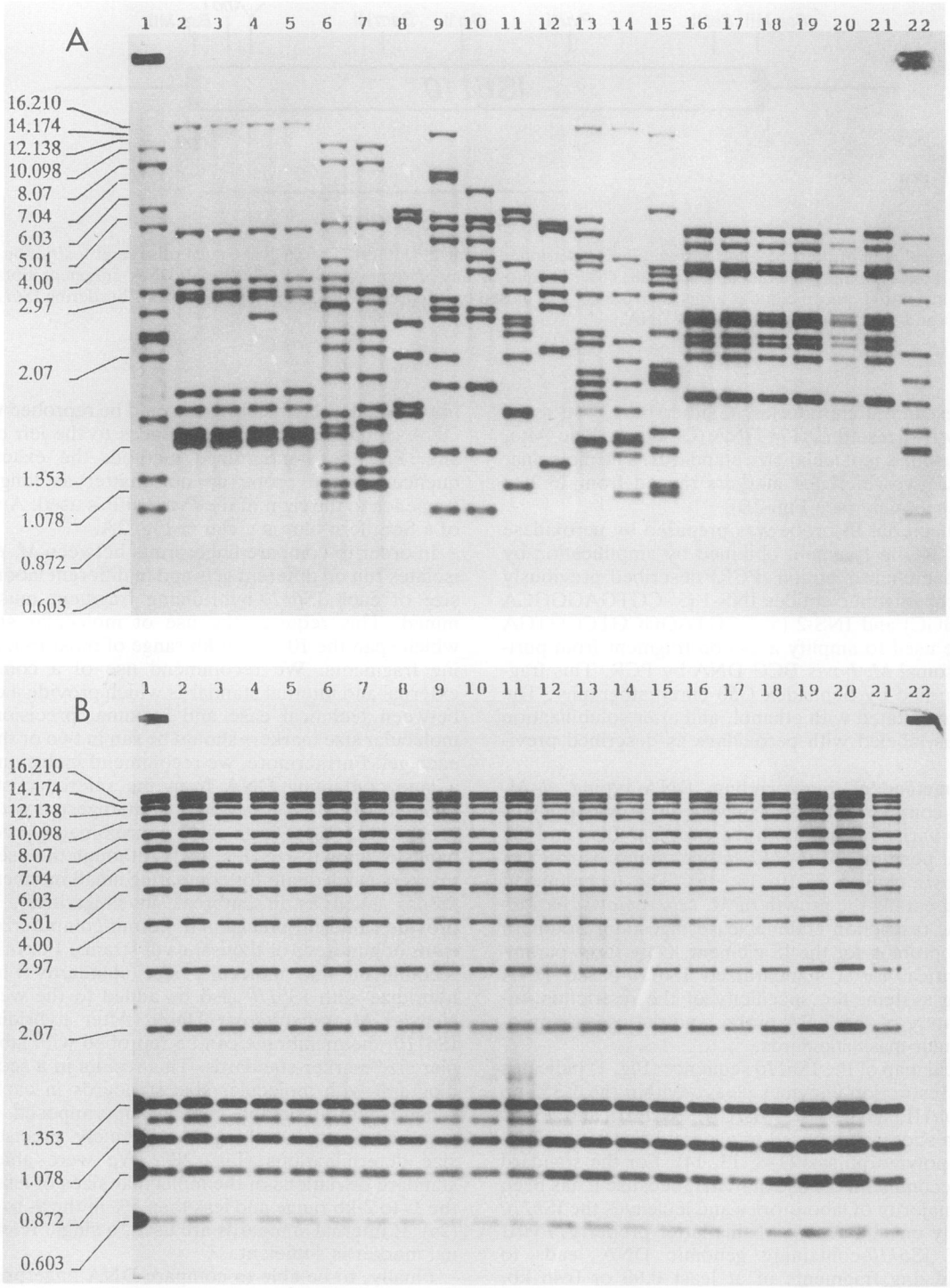

FIG. 2. Fingerprints of *M. tuberculosis* strains obtained by the recommended standard method. Chromosomal DNAs of 22 different mycobacterial isolates were digested with *Pvu*II, and after mixing with marker DNA, the fragments were separated by overnight electrophoresis. The fragments were transferred to filters and hybridized with peroxidase-labeled IS*6110* DNA (A) and peroxidase-labeled marker DNA (B). Lane 22, reference *M. tuberculosis* Mt14323, which is available on request; lanes 2 to 5, strains from an outbreak of tuberculosis in Amsterdam during 1991. All other lanes contain DNAs from *M. tuberculosis* isolated from epidemiologically unrelated cases. The strains corresponding to lanes 16 to 21 were selected from a collection of about 200 randomly chosen strains from patients in The Netherlands. The internal DNA markers (B) were *Pvu*II-digested supercoiled ladder DNAs with molecular sizes of 16.2, 14.2, 12.1, 10.1, 8.07, 7.04, 6.03, 5.01, 4.00, 2.97, and 2.07 kb and *Hae*III-digested ϕX174 with molecular sizes of 1,353, 1,078, 872, and 603 bp, respectively. A detailed protocol on the procedures for DNA isolation and fingerprinting is available from one of us (J.D.A.V.).

electrophoresis, blotting, and hybridization, provided that the resolving power of the electrophoresis procedure is within the DNA fragment size range of 0.9 to 10 kb.

Dick van Soolingen and Petra de Haas are acknowledged for ideas and technical assistance. Janetta Top and Wilma van der Roest are acknowledged for secretarial help.

This study was financially supported by the World Health Organization Program for Vaccine Development and the European Community Program on Science Technology and Development.

REFERENCES

1. **Cave, M. D., K. D. Eisenach, P. F. McDermott, J. H. Bates, and J. T. Crawford.** 1991. IS*6110*: conservation of sequence in the *Mycobacterium tuberculosis* complex and its utilization in DNA fingerprinting. Mol. Cell. Probes **5**:73–80.
2. **Daley, C. L., P. M. Small, G. F. Schecter, G. K. Schoolnik, R. A. McAdam, W. R. Jacobs, Jr., and P. C. Hopewell.** 1992. An outbreak of tuberculosis with accelerated progression among persons infected with the human immunodeficiency virus: an analysis using restriction fragment length polymorphisms. N. Engl. J. Med. **326**:231–235.

2a. **de Haas, P.** Unpublished data.

3. **Edlin, B. R., J. I. Tokars, M. H. Grieco, J. T. Crawford, J. Williams, E. M. Sordillo, K. R. Ong, J. O. Kilburn, S. W. Dooley, K. G. Castro, W. R. Jarvis, and S. D. Holmberg.** 1992. An outbreak of multidrug-resistant tuberculosis among hospitalized patients with the acquired immunodeficiency syndrome. N. Engl. J. Med. **326**:1514–1521.
4. **Eisenach, K. D., J. T. Crawford, and J. H. Bates.** 1988. Repetitive DNA sequences as probes for *Mycobacterium tuberculosis*. J. Clin. Microbiol. **26**:2240–2245.
5. **Fomukong, N. G., J. W. Dale, T. W. Osborn, and J. M. Grange.** 1992. Use of gene probes based on the insertion sequence IS*986* to differentiate between BCG vaccine strains. J. Appl. Microbiol. **72**:126–133.
6. **Hermans, P. W. M., D. van Soolingen, E. M. Bik, P. E. W. de Haas, J. W. Dale, and J. D. A. van Embden.** 1991. The insertion element IS*987* from *M. bovis* BCG is located in a hot spot integration region for insertion elements in *M. tuberculosis* complex strains. Infect. Immun. **59**:2695–2705.
7. **Hermans, P. W. M., D. van Soolingen, J. W. Dale, A. R. Schuitema, R. A. McAdam, D. Catty, and J. D. A. van Embden.** 1990. Insertion element IS*986* from *Mycobacterium tuberculosis*: a useful tool for diagnosis and epidemiology of tuberculosis. J. Clin. Microbiol. **28**:2051–2058.
8. **Mazurek, G. H., M. D. Cave, K. D. Eisenach, R. J. Wallace, Jr., J. H. Bates, and J. T. Crawford.** 1991. Chromosomal DNA fingerprint patterns produced with IS*6110* as strain specific markers for epidemiologic study of tuberculosis. J. Clin. Microbiol. **29**:2030–2033.
9. **McAdam, R. A., P. W. M. Hermans, D. van Soolingen, Z. F. Zainuddin, D. Catty, J. D. A. van Embden, and J. W. Dale.** 1990. Characterization of a *Mycobacterium tuberculosis* insertion sequence belonging to the IS*3* family. Mol. Microbiol. **4**:1607–1613.
10. **Otal, I., C. Martin, V. Vincent-Lévy-Frébault, D. Thierry, and B. Gicquel.** 1991. Restriction fragment length polymorphism analysis using IS*6110* as an epidemiological marker in tuberculosis. J. Clin. Microbiol. **29**:1252–1254.
11. **Thierry, D., A. Brisson-Noël, V. Vincent-Lévy-Frébault, S. Nguyen, J. Guesdon, and B. Gicquel.** 1990. Characterization of a *Mycobacterium tuberculosis* insertion sequence, IS*6110*, and its application in diagnosis. J. Clin. Microbiol. **28**:2668–2673.
12. **Thierry, D., M. D. Cave, K. D. Eisenach, J. T. Crawford, J. H. Bates, B. Gicquel, and J. L. Guesdon.** 1990. IS*6110*, an IS-like element of *Mycobacterium tuberculosis* complex. Nucleic Acids Res. **18**:188.
13. **Van Soolingen, D., P. W. M. Hermans, P. E. W. de Haas, D. R. Soll, and J. D. A. van Embden.** 1991. The occurrence and stability of insertion sequences in *Mycobacterium tuberculosis* complex strains; evaluation of IS-dependent DNA polymorphism as a tool in the epidemiology of tuberculosis. J. Clin. Microbiol. **29**:2578–2586.
14. **Zainuddin, Z. F.** 1988. Mycobacterial plasmids and related DNA sequences. Ph.D. thesis. University of Surrey, Surrey, United Kingdom.
15. **Zainuddin, Z. F., and J. W. Dale.** 1989. Polymorphic repetitive DNA sequences in *Mycobacterium tuberculosis* detected with a gene probe from a *Mycobacterium fortuitum* plasmid. J. Gen. Microbiol. **135**:2347–2355.